PATENT APPLICATION
ATTORNEY DOCKET NO. 10122.009001

APPLICATION

FOR

UNITED STATES LETTERS PATENT

TITLE: **MONITORING APPARATUS**

APPLICANTS: **Takahiro HAMADA; Takuma KAWAMURA
Fuminori WATANABE and John Alexenko**

"EXPRESS MAIL" Mailing Label Number: EV804240065US
Date of Deposit: JULY 28, 2006

DESCRIPTION

MONITORING APPARATUS

Technical Field

The present invention relates to a monitoring apparatus, and more particularly, to a monitoring apparatus suitable for monitoring a digital video/audio signal.

Background Art

In recent years, video processing techniques have been improved, and thus a high-quality video broadcast, such as a high-definition television broadcast, has been provided. Here, digital video signals of a high-definition television broadcast, etc., are often transmitted to each home through satellite broadcasting and a cable TV network. However, an error sometimes occurs during the transmission of video signals from various causes. When an error occurs, problems, such as a video freeze, a blackout, noise, audio mute, etc., may result, and thus it becomes necessary to take countermeasures.

Under these circumstances, in Japanese Patent Application Laid-Open No. 2003-204562, the present applicant discloses, for example a signal monitoring system in which

a central processing terminal calculates the difference between a first statistic value based on a video signal (first signal) output from a transmission source and a second statistic value based on a video signal (second signal) output from a relay station or a transmission destination, if the difference is below a threshold value, the transmission is determined to be normal, and if the difference is over the threshold value, a determination is made that a transmission trouble has occurred between the transmission source 10 and the relay station 20 so that a warning signal is output to raise an alarm (alarm display and alarm sound).

However, in the case of such a signal monitoring system, the central processing terminal only calculates the difference between the first statistic value and the second statistic value, and makes an automatic determination of a transmission trouble on the basis of the difference. Thus, there is a problem in that a sufficient analysis cannot be made on what kind of error has occurred from that determination. If a sufficient error analysis cannot be made, the same trouble may be brought about again. Here, in order to analyze an error, it is necessary to store all the transmitted video signals, and to check the signals while reproducing the signals later with taking time. However, in order to do that, it needs a vast amount of storage capacity for storing video/audio signals and enormous checking time.

Disclosure of Invention

The present invention has been made in view of the problems of these known techniques. It is an object of the present invention to provide a monitoring system capable of analyzing the contents of the error if an error occurs.

According to the present invention, there is provided a monitoring system for monitoring a video/audio signal transmitted from a transmission source to a transmission destination, the system including: a step of storing the video/audio signal transmitted from the transmission source to the transmission destination repeatedly for a predetermined time period; a step of comparing a first characteristic amount extracted from the video/audio signal before the transmission and a second characteristic amount extracted from the video/audio signal after the transmission in real time; a step of determining an error occurrence when there is a difference of a predetermined value or more between the first characteristic amount and the second characteristic amount; and a step of transmitting the stored video/audio signal to a predetermined destination when an error occurrence is determined.

By the monitoring system of the present invention, a video/audio signal transmitted from a transmission source

to a transmission destination is stored repeatedly for a predetermined time period. Thus, for example if the video/audio signal is overwritten at predetermined timing, it is sufficient to have a small storage capacity for the video/audio signal. Also, the video/audio signal on which an error has occurred remains without being overwritten at the time of determination that an error has occurred. Accordingly, by transmitting this signal to a predetermined destination, a detailed error analysis can be made promptly on the basis of the transmitted video/audio signal. In this regard, a "video/audio signal" generally refers to a signal including both a picture signal (video signal) and a sound signal (audio signal). However, in this specification, a signal including at least one of those signals is sufficient, and it does not matter whether the signal is raw data or compressed data.

In the monitoring system, the second characteristic amount to be used for comparison and the stored video/audio signal are preferably transmitted from the transmission destination to the transmission source through the Internet.

In the monitoring system, the stored video/audio signal is preferably used for analyzing the error.

In the monitoring system, the error is preferably an

image freeze phenomenon.

In the monitoring system, the error is preferably a blackout phenomenon.

In the monitoring system, the error is preferably an audio mute phenomenon.

In the monitoring system, the error is preferably an audio failure phenomenon.

In the monitoring system, the error is preferably a video/audio mismatching phenomenon.

In the monitoring system, the error is preferably an irregular frame phenomenon.

In the monitoring system, the video/audio signal transmitted to the transmission destination is preferably corrected when there is a difference of a predetermined value or more between the first characteristic amount and the second characteristic amount.

Brief Description of the Drawings

Fig. 1 is a schematic diagram of the entire transmission

system including a monitoring system according to the present embodiment.

Fig. 2 is a block diagram illustrating the configuration of extraction apparatuses 100X, 100A, and 100B.

Fig. 3A is a diagram showing the characteristic amount extracted by the extraction apparatus 100X. Fig. 3C is a diagram showing the characteristic amount extracted by the extraction apparatus 100A or 100B. Fig. 3B is a diagram showing the value of a difference between the two characteristic amounts.

Fig. 4A is a diagram showing the characteristic amount extracted by the extraction apparatus 100X. Fig. 4C is a diagram showing the characteristic amount extracted by the extraction apparatus 100A or 100B. Fig. 4B is a diagram showing the value of a difference between the two characteristic amounts.

Fig. 5A is a diagram showing the characteristic amount extracted by the extraction apparatus 100X. Fig. 5C is a diagram showing the characteristic amount extracted by the extraction apparatus 100A or 100B. Fig. 5B is a diagram showing the value of a difference between the two characteristic amounts.

Fig. 6A is a diagram showing the characteristic amount extracted by the extraction apparatus 100X. Fig. 6C is a diagram showing the characteristic amount extracted by the extraction apparatus 100A or 100B. Fig. 6B is a diagram showing the value of a difference between the two characteristic amounts.

Fig. 7A is a diagram showing the characteristic amount extracted by the extraction apparatus 100X. Fig. 7C is a diagram showing the characteristic amount extracted by the extraction apparatus 100A or 100B. Fig. 7B is a diagram showing the value of a difference between the two characteristic amounts.

Fig. 8A is a diagram showing the characteristic amount extracted by the extraction apparatus 100X. Fig. 8C is a diagram showing the characteristic amount extracted by the extraction apparatus 100A or 100B. Fig. 8B is a diagram showing the value of a difference between the two characteristic amounts.

Best Mode for Carrying out the Invention

In the following, a description will be given of the present invention with reference to an embodiment. Fig. 1

is a schematic diagram of the entire transmission system including a monitoring system according to the present embodiment. In Fig. 1, consider the case where for example, a video/audio signal including an audio signal and a video signal is transmitted from a transmission source 10, such as a broadcasting station to transmission destinations 20A and 20B, such as satellite stations. An example in which the transmission of such a video/audio signal is carried out through a communication satellite S is shown. However, the transmission may be through various means, for example via optical fibers.

Fig. 2 is a block diagram illustrating the configuration of extraction apparatuses 100X, 100A, and 100B. Among the video/audio signals, left and right audio signals AL and AR are input into audio input sections 101 and 102, respectively. The signals output from there are input into delay sections 103 and 104, respectively. The operation result of an audio section 105 is output as audio characteristic amount (Audio Level, Audio Activity), and thus is output from the extraction apparatuses 100X, 100A, and 100B to the terminals 200X, 200A, and 200B. Here, the Audio Level means an average value of the absolute values of the audio sampling (48 KHz) values (48000/30 = 1600 pieces) included in one frame (for example, 30 frames/s) of an image. Also, the Audio Activity means a square mean value of the

audio sampling (48 KHz) values (48000/30 = 1600 pieces) included in one frame (for example, 30 frames/s).

On the other hand, among the video/audio signals, a video signal VD is input into a video input section 108. The signal output from there is input into frame memories 109, 110, and 111. The frame memory 109 stores the current frame, the frame memory 110 stores the previous frame, and the frame memory 111 stores the frame before two.

The output signals from the frame memories 109, 110, and 111 are input into an MC calculation section 112, and the calculation result thereof is output as the characteristic amount (Motion) of the video. At the same time, the output signal from the frame memory 110 is input into a video calculation section 119. The calculation result of the video calculation section 119 is output as the characteristic amount (Video Level, Video Activity) of the video. These output signals are output from the extraction apparatuses 100X, 100A, and 100B to the terminals 200X, 200A, and 200B. Here, the Motion is calculated as follows. For example, an image frame is divided into 8 pixels × 8 line-size small blocks, the average value and the variance of the 64 pixels are calculated for each small block, and the Motion is represented by the difference between the average value and the variance of the blocks of the same place of the frame

before N, and indicates the movement of the image. Note that N is normally any one of 1, 2, and 4. Also, the Video Level is the average value of the pixel values included in an image frame. Furthermore, for the Video Activity, when a variance is obtained for each small block included in an image, the average value of the variances of the pixels included in a frame may be used. Alternatively, the variance of the pixels in a frame included in an image frame may simply be used.

Next, a description will be given of the operation of this monitoring system. The video/audio signal before being transmitted from the transmission source 10 is input into the extraction apparatus 100X extracting a characteristic amount (meta data) therefrom, and is stored into the video server 201X while being overwritten for a predetermined time period.

On the other hand, the previous video/audio signal having been transmitted to the transmission destinations 20A and 20B is input into the extraction apparatuses 100A and 100B which extract characteristic amounts therefrom, respectively, and is stored into the video servers 201A and 201B while being overwritten for a predetermined time period (a step of storing the video/audio signal transmitted from the transmission source to the transmission destination

repeatedly for a predetermined time period). The extracted characteristic amount is transmitted from the terminals 200A and 200B to the terminal 200X of the transmission source 10 through the Internet INT.

The terminal 200X of the transmission source 10 compares, in real time, the characteristic amount extracted by the extraction apparatus 100X and the characteristic amounts transmitted from the terminals 200A and 200B (a step of comparing, in real time, the first characteristic amount extracted from the video/audio signal before transmission and the second characteristic amount extracted from the video/audio signal after transmission). If there is a difference of a predetermined value or more, a determination is made that an error has occurred (a step of determining that an error has occurred if there is a difference of a predetermined value or more between the first characteristic amount and the second characteristic amount), the terminal 200X transmits an instruction to the terminals 200A and 200B through the Internet INT. The terminals 200A and 200B clip the video/audio signal of the past 10 seconds from the current point in time from the video servers 201A and 201B in response to this instruction, and transmit the video clip to the terminal 200X of the transmission source 10 through the Internet INT (a step of transmitting the stored video/audio signal to a predetermined destination when a determination

is made that an error has occurred).

An operator of the transmission source 10 can compare the video/audio signal transmitted from the terminals 200A and 200B to the terminal 200X and the video/audio signal stored in the video server 201X, and analyze the cause of the error. In this regard, the terminal 200X may transmit the characteristic amount of the video/audio signal before the transmission to the terminals 200A and 200B through the Internet INT, and the transmission destinations may compare the characteristic amounts.

Next, a more detailed description will be given of errors.

(1) Detection of image freeze phenomenon

Fig. 3A is a diagram showing the characteristic amount extracted by the extraction apparatus 100X. Fig. 3C is a diagram showing the characteristic amount extracted by the extraction apparatus 100A or 100B. Fig. 3B is a diagram showing the value of a difference between the two characteristic amounts. The vertical axis represents a motion [value] as a characteristic amount, and the horizontal axis represents time.

Here, as shown in Fig. 3C, the motion [value] is low during the time t1 to t2 in the video based on the video/audio

signal after the transmission. As shown in Fig. 3A, the motion [value] is also low during the time t1 to t2 in the video based on the video/audio signal before the transmission, and thus the difference between the two is zero (refer to Fig. 3B). This has occurred, because the transmitted picture is a still image. Accordingly, it can be determined that an image freeze phenomenon has not occurred.

On the other hand, as shown in Fig. 3C, the motion [value] is low during the time t3 to t4 in the video based on the video/audio signal after the transmission. However, as shown in Fig. 3A, the motion [value] is high during the time t3 to t4 in the video based on the video/audio signal before the transmission, and thus the difference between the two exceeds a threshold value TH1 (refer to Fig. 3B). This is due to the fact that an image freeze phenomenon has occurred because of some cause. The terminal 200X of the transmission source 10, which has detected this fact, transmits an instruction to transmit the video clip immediately to the transmission destination terminal 200A or 200B, in which the problem has occurred.

(2) Detection of blackout phenomenon

Fig. 4A is a diagram showing the characteristic amount extracted by the extraction apparatus 100X. Fig. 4C is a diagram showing the characteristic amount extracted by the

extraction apparatus 100A or 100B. Fig. 4B is a diagram showing the value of a difference between the two characteristic amounts. The vertical axis represents a Video Activity [value] as a characteristic amount, and the horizontal axis represents time. For this Video Activity [value], for example the following variance A can be used.

Considering the video signals (assuming the case of having values for individual three-dimensional coordinate values such as a virtual video signal. Note that the signals become two-dimensional video signals assuming that z = 0) before and after the transmission, it is assumed that the video signal at the three-dimensional coordinate (x, y, z) at time t before the transmission is V (x, y, z, t), and the video signal at the three-dimensional coordinate (x, y, z) at time t after the transmission is U (x, y, z, t).

Here, if a video signal is transmitted over a long distance, various problems, such as a signal loss, noise, etc., might occur, and thus it is not always true that V (x, y, z, t) = U (x, y, z, t). It is not necessary to correct an error if viewers are not aware of the error. However, problems, such as a blackout phenomenon, need the countermeasures.

The variance A as the characteristic amount of the

video signal V (x, y, z, t) can be represented by the following expression.

$$\text{Variance } A = \frac{1}{XYZT}\sum_{x=1}^{X}\sum_{y=1}^{Y}\sum_{z=1}^{Z}\sum_{t=1}^{T}\{V(x,y,z,t) - ave.V\}^2$$

Also, the average value ave.V can be obtained by the following expression.

$$\text{Average } ave.V = \frac{1}{XYZT}\sum_{x=1}^{X}\sum_{y=1}^{Y}\sum_{z=1}^{Z}\sum_{t=1}^{T}V(x,y,z,t)$$

The variance A is obtained for each of the video signal V (x, y, z, t) before the transmission and the video signal U (x, y, z, t) after the transmission. Then the difference between the two is obtained, and thus a blackout is determined as follows.

As shown in Fig. 4C, the variance value is low during the time t1 to t2 in the video based on the video/audio signal after the transmission. As shown in Fig. 4A, the variance value is also low during the time t1 to t2 in the video based on the video/audio signal before the transmission, and thus the difference between the two is zero (refer to Fig. 4B). This has occurred, because the transmitted picture is, for example a picture of a starry sky. Accordingly, it can be

determined that an image freeze phenomenon has not occurred.

On the other hand, as shown in Fig. 4C, the variance value is low during the time t3 to t4 in the video based on the video/audio signal after the transmission. However, as shown in Fig. 4A, the variance value is high during the time t3 to t4 in the video based on the video/audio signal before the transmission, and thus the difference between the two exceeds a threshold value TH2 (refer to Fig. 4B). This is due to the fact that a blackout phenomenon of having an all-black screen has occurred because of some cause. The terminal 200X of the transmission source 10, which has detected this fact, transmits an instruction to transmit the video clip immediately to the transmission destination terminal 200A or 200B, in which the problem has occurred.

(3) Detection of audio mute phenomenon

Fig. 5A is a diagram showing the characteristic amount extracted by the extraction apparatus 100X. Fig. 5C is a diagram showing the characteristic amount extracted by the extraction apparatus 100A or 100B. Fig. 5B is a diagram showing the value of a difference between the two characteristic amounts. The vertical axis represents an Audio Level [value] as a characteristic amount, and the horizontal axis represents time. In this regard, the sampling of the Audio Level [value] of the audio signal is

preferably leveled off by the frame frequency of the video signal. For example, in the case of a video signal having 30 frames per one second, the sampling of the Audio Level [value] is preferably performed by 30 Hz.

Here, as shown in Fig. 5C, the Audio Level [value] is very low during the time t1 to t2 in the audio based on the video/audio signal after the transmission. As shown in Fig. 5A, the Audio Level [value] is also low during the time t1 to t2 in the audio based on the video/audio signal before the transmission, and thus the difference between the two is zero (refer to Fig. 5B). This has occurred, because the Audio Level [value] of the video/audio signal before transmission is originally low. Accordingly, it can be determined that an audio mute phenomenon has not occurred.

On the other hand, as shown in Fig. 5C, the Audio Level [value] is low during the time t3 to t4 in the audio based on the video/audio signal after the transmission. However, as shown in Fig. 5A, the Audio Level [value] is high during the time t3 to t4 in the audio based on the video/audio signal before the transmission, and thus the difference between the two exceeds a threshold value TH3 (refer to Fig. 5B). This is due to the fact that an audio mute phenomenon has occurred because of some cause. The terminal 200X of the transmission source 10, which has detected this fact,

transmits an instruction to transmit the video clip immediately to the transmission destination terminal 200A or 200B, in which the problem has occurred.

(4) Detection of audio failure phenomenon

Fig. 6A is a diagram showing the characteristic amount extracted by the extraction apparatus 100X. Fig. 6C is a diagram showing the characteristic amount extracted by the extraction apparatus 100A or 100B. Fig. 6B is a diagram showing the value of a difference between the two characteristic amounts. The vertical axis represents an Audio Level [value] as a characteristic amount, and the horizontal axis represents time. In this regard, the sampling of the Audio Level [value] of the audio signal is preferably leveled off by the frame frequency of the video signal.

Here, when the difference between the Audio Level [value] based on the video/audio signal after the transmission and the Audio Level [value] based on the video/audio signal before the transmission is obtained, as shown in Fig. 6B, the difference between the two exceeds a threshold value TH4 during the time t1 to t2 and the time t3 to t4. This is due to the fact that an audio failure phenomenon has occurred by an overlap of noise, etc., because of some cause. The terminal 200X of the transmission source

10, which has detected this fact, transmits an instruction to transmit the video clip immediately to the transmission destination terminal 200A or 200B, in which the problem has occurred.

(5) Detection of video/audio mismatching phenomenon

Fig. 7A is a diagram showing the Audio Level [value] extracted by the extraction apparatus 100X in response to a video frame. Fig. 7C is a diagram showing the Audio Level [value] extracted by the extraction apparatus 100A or 100B. Fig. 7B is a diagram showing an audio advance/delay regarding time. In this regard, the sampling of the Audio Level [value] of the audio signal is preferably leveled off by the frame frequency of the video signal.

Here, rises of the Audio Level [value] with respect to a frame are detected, and compared. As shown in Fig. 7B, the audio delay amount with respect to video exceeds a threshold value TH5+ at the time t1 and t3, and the audio advance amount with respect to video is below a threshold value TH5- at the time t2. The terminal 200X of the transmission source 10, which has determined that a video/audio mismatching phenomenon has occurred by detecting either of the two, transmits an instruction to transmit the video clip immediately to the transmission destination terminal 200A or 200B, in which the problem has occurred.

(6) Detection of irregular frame phenomenon

Fig. 8A is a diagram showing the characteristic amount extracted by the extraction apparatus 100X. Fig. 8C is a diagram showing the characteristic amount extracted by the extraction apparatus 100A or 100B. Fig. 8B is a diagram showing the value of a difference between the two characteristic amounts. The vertical axis represents a Video Level [value] (the above-described variance may be used) as a characteristic amount, and the horizontal axis represents time.

When the video based on the video/audio signal before the transmission completely matches the video based on the video/audio signal after the transmission, the difference of the statistic values of the image values becomes zero. However, if a video signal of one frame is inserted into the video/audio signal after the transmission, the difference exceeds a predetermined threshold value.

As shown in Fig. 8B, the difference of the statistic values of the pixel values exceeds a threshold value TH6+ during the time t1 to t2, and the difference of the statistic values of the pixel values is below a threshold value TH6- during the time t3 to t4. The terminal 200X of the transmission source 10, which has determined that an

irregular frame phenomenon has occurred by detecting either of the two, transmits an instruction to transmit the video clip immediately to the transmission destination terminal 200A or 200B, in which the problem has occurred.

In this regard, after the transmission source 10 has analyzed an error, for example, if an irregular frame phenomenon occurs, the terminal 200X of the transmission source can transmit an instruction to eliminate the misalignment between the video signal and the audio signal to the transmission destination terminal 200A or 200B. The transmission destination terminal 200A or 200B can perform the processing on the video/audio signal to eliminate the misalignment between the video signal and the audio signal. Such an instruction can be transmitted to the transmission destination terminals all together using the Internet.